STAT 713 hw 2

Maximum likelihood and method of moments estimators

Do problems 7.1, 7.6, 7.10, 7.13, 7.14, 7.15(a) from CB. In addition:

1. For the following, find (i) MLEs and (ii) method of moment estimators for the unknown parameters:

   (a) $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Uniform}(a, b)$, $-\infty < a < b < \infty$.

   The find the method of moments estimator we can use the equations

$$
\begin{aligned}
m_1 &= \frac{a+b}{2} \\
m_2 &= \int_a^b \frac{x^2}{b-a} dx = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b}{3},
\end{aligned}
$$

   or, equivalently, we can use

$$
\begin{aligned}
m_1 &= \frac{a+b}{2} \\
m_2 - m_1^2 &= \frac{(b-a)^2}{12}.
\end{aligned}
$$

   These both give

$$
\begin{aligned}
a &= m_1/2 - \sqrt{3(m_2 - m_1^2)} \\
b &= m_1/2 + \sqrt{3(m_2 - m_1^2)}.
\end{aligned}
$$

   so the method of moments estimators for a and b are

$$
\begin{aligned}
\bar{a} &= \hat{m}_1/2 - \sqrt{3(\hat{m}_2 - \hat{m}_1^2)} \\
\bar{b} &= \hat{m}_1/2 + \sqrt{3(\hat{m}_2 - \hat{m}_1^2)},
\end{aligned}
$$

   where $\hat{m}_1 = n^{-1} \sum_{i=1}^n X_i$ and $\hat{m}_2 = n^{-1} \sum_{i=1}^n X_i^2$.

   The MLEs for a and b are $\hat{a} = X_{(1)}$ and $\hat{b} = X_{(n)}$.

   (b) $X_1, \dots, X_n \overset{\text{ind}}{\sim} f(x; \mu, \sigma) = (1/\sigma) e^{-(x-\mu)/\sigma} \cdot \mathbf{1}(x > \mu)$, $\mu \in \mathbb{R}$, $\sigma > 0$.

   We note that for all $\mu \in \mathbb{R}$, $\sigma > 0$, and $x \in \mathbb{R}$, we may write $f(x; \mu, \sigma) = (1/\sigma) f_Z((x-\mu)/\sigma)$, where $f_Z(z) = e^{-z} \cdot \mathbf{1}(z > 0)$. So if $X \sim f(x; \mu, \sigma)$, we can write $X = \mu + \sigma Z$, where $Z \sim f_Z$. Noting that $\mathbb{E}Z = 1$ and $\text{Var}\, Z = 1$, we may write

$$
\begin{aligned}
m_1 &= \mathbb{E}X = \mathbb{E}(\mu + \sigma Z) = \mu + \sigma \\
m_2 - m_1^2 &= \text{Var}\, X = \text{Var}(\mu + \sigma Z) = \sigma^2.
\end{aligned}
$$

Solving for μ and σ gives

$$\sigma = \sqrt{m_2 - m_1^2}$$
$$\mu = m_1 - \sqrt{m_2 - m_1^2}$$

so the method of moments estimators for σ and μ are given by

$$\bar{\sigma} = \sqrt{\hat{m}_2 - \hat{m}_1^2}$$
$$\bar{\mu} = \hat{m}_1 - \sqrt{\hat{m}_2 - \hat{m}_1^2},$$

where $\hat{m}_1 = n^{-1}\sum_{i=1}^n X_i$ and $\hat{m}_2 = n^{-1}\sum_{i=1}^n X_i^2$.

We find the MLEs as follows: The likelihood function is given by

$$\mathcal{L}(\sigma, \mu; \mathbf{X}) = (1/\sigma)^n e^{-n\bar{X}_n/\sigma} e^{n\mu} \mathbf{1}(X_{(1)} > \mu).$$

We see that for each $\sigma > 0$, $\mathcal{L}(\sigma, \mu; \mathbf{X})$ is an increasing function of μ when $\mu \leq X_{(1)}$. So the MLE for μ is $\hat{\mu} = X_{(1)}$. Now the MLE for σ is the maximizer of $\mathcal{L}(\sigma, X_{(1)}; \mathbf{X})$. We find this by maximizing

$$\log \mathcal{L}(\sigma, X_{(1)}; \mathbf{X}) = -n \log \sigma - n(\bar{X}_n - X_{(1)})/\sigma.$$

We have

$$\frac{\partial}{\partial \sigma} \log \mathcal{L}(\sigma, X_{(1)}; \mathbf{X}) = -\frac{n}{\sigma} + \frac{n(\bar{X}_n - X_{(1)})}{\sigma^2}.$$

Setting this equal to zero and solving for σ gives

$$\hat{\sigma} = \bar{X}_n - X_{(1)}.$$

Problems 7.1, 7.6, 7.10, 7.13, 7.14, 7.15 (a)

$\boxed{7.1}$ The MLE $\hat{\theta}_n$ is the value of θ which maximizes $f(x;\theta)$.

So for the setup

x	$f(x\|1)$	$f(x\|2)$	$f(x\|3)$
0	$\frac{1}{3}$	$\frac{1}{4}$	0
1	$\frac{1}{3}$	$\frac{1}{4}$	0
2	0	$\frac{1}{4}$	$\frac{1}{4}$
3	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{2}$
4	$\frac{1}{6}$	0	$\frac{1}{4}$

we have

$$\hat{\theta}_n \in \begin{cases} \{1\} & \text{for } x \in \{0,1\} \\ \{2,3\} & \text{for } x = 2 \\ \{3\} & \text{for } x \in \{3,4\} \end{cases}$$

$\longleftarrow$ No unique maximizer

$\boxed{7.6}$ Let $X_1,\dots,X_n \overset{ind}{\sim} f(x;\theta) = \theta x^{-2} \mathbb{1}(x \geq \theta)$.

(a) We have $f(\underset{\sim}{x};\theta) = \underbrace{\theta^n \mathbb{1}(x_{(1)} \geq \theta)}_{g(x_{(1)};\theta)} \Big(\underbrace{\prod_{i=1}^n}_{h(\underset{\sim}{x})} x_i\Big)^{-2}$,

so $T(\underset{\sim}{X}) = X_{(1)}$ is a suff. statistic.

(b) The MLE for θ is

$$\hat{\theta}_n = \underset{\theta}{\arg\max}\ L(\theta;\underset{\sim}{X})$$

$$= \underset{\theta}{\arg\max}\ \theta^n \mathbb{1}(X_{(1)} \geq \theta)$$

$$= X_{(1)}.$$

(c) The MoM estimator does not exist because $m_1 = \int_\theta^\infty x \cdot \theta x^{-2}\,dx$ does not converge.

$\boxed{7.10}$ Let $X_1, \dots, X_n \overset{ind}{\sim} F_X(x; \alpha, \beta) = (x/\beta)^\alpha$ for $x \in [0, \beta]$, $\alpha, \beta > 0$.

(a) We have $f_X(x; \alpha, \beta) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \mathbb{1}(0 \le x \le \beta)$ and

$$f(\underset{\sim}{x}; \alpha, \beta) = \underbrace{\left(\frac{\alpha}{\beta}\right)^n \left(\prod_{i=1}^n x_i\right)^{\alpha-1} \left(\frac{1}{\beta}\right)^{n(\alpha-1)} \mathbb{1}(x_{(n)} \le \beta)}_{g\left(\prod_{i=1}^n x_i,\ x_{(n)}\right)} \underbrace{\mathbb{1}(x_{(1)} \ge 0)}_{h(\underset{\sim}{x})},$$

so $T(\underset{\sim}{X}) = \left(\prod_{i=1}^n X_i,\ X_{(n)}\right)$ is a sufficient statistic.

(b) The likelihood function is given by

$$L(\alpha, \beta; \underset{\sim}{X}) = \left(\frac{\alpha}{\beta}\right)^n \left(\prod_{i=1}^n X_i\right)^{\alpha-1} \left(\frac{1}{\beta}\right)^{n(\alpha-1)} \mathbb{1}(X_{(n)} \le \beta),$$

which, for any α, is maximized in β at $\beta = X_{(n)}$.
Plugging in $\beta = X_{(n)}$ we have

$$L(\alpha, X_{(n)}; \underset{\sim}{X}) = \left(\frac{\alpha}{X_{(n)}}\right)^n \left(\prod_{i=1}^n X_i\right)^{\alpha-1} \left(\frac{1}{X_{(n)}}\right)^{n(\alpha-1)}$$

and

$$\ell(\alpha, X_{(n)}; \underset{\sim}{X}) = n \log \alpha - n \log X_{(n)} + (\alpha-1)\sum_{i=1}^n \log X_i - n(\alpha-1)\log X_{(n)}$$

$$= n \log \alpha - n\alpha \log X_{(n)} + (\alpha-1)\sum_{i=1}^n \log X_i .$$

Now

$$\frac{\partial}{\partial\alpha}\ell(\alpha, X_{(n)}; \underset{\sim}{X}) = \frac{n}{\alpha} - n \log X_{(n)} + \sum_{i=1}^n \log X_i \overset{set}{=} 0$$

$$\iff \alpha = \frac{n}{n \log X_{(n)} - \sum_{i=1}^n \log X_i}$$

So the MLEs are $(\hat\alpha, \hat\beta) = \left(\frac{n}{n \log X_{(n)} - \sum_{i=1}^n \log X_i},\ X_{(n)}\right)$.

$\boxed{7.13}$ Let $X_1, \dots, X_n \overset{iid}{\sim} f(x;\theta) = \frac{1}{2} e^{-|x-\theta|}$, $\theta \in \mathbb{R}$.

The likelihood function is given by

$$L(\theta; \underset{\sim}{X}) = \left(\frac{1}{2}\right)^n \prod_{i=1}^{n} e^{-|x_i - \theta|}$$

$$= \left(\frac{1}{2}\right)^n \exp\left[-\sum_{i=1}^{n} |x_{(i)} - \theta|\right].$$

Note that $\hat{\theta}_n = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^{n} |x_{(i)} - \theta|$.

For $\theta \in [x_{(j)}, x_{(j+1)}]$, we have

$$\sum_{i=1}^{n} |x_{(i)} - \theta| = \begin{cases} \sum_{i=1}^{n} (x_{(i)} - \theta) & \theta < x_{(1)} \\ \sum_{i=2}^{n} (x_{(i)} - \theta) - (x_{(1)} - \theta) & x_{(1)} \le \theta < x_{(2)} \\ \vdots & \vdots \\ \sum_{i=j+1}^{n} (x_{(i)} - \theta) - \sum_{i=1}^{j} (x_{(i)} - \theta) & x_{(j)} \le \theta < x_{(j+1)} \\ \vdots & \vdots \\ (x_{(n)} - \theta) - \sum_{i=1}^{n-1} (x_{(i)} - \theta) & x_{(n-1)} \le \theta < x_{(n)} \\ -\sum_{i=1}^{n} (x_{(i)} - \theta) & x_{(n)} \le \theta \end{cases}$$

$$= \begin{cases} \sum_{i=1}^{n} x_{(i)} - n\theta & \theta < x_{(1)} \\ \sum_{i=2}^{n} x_{(i)} - x_{(1)} + (2-n)\theta & x_{(1)} \le \theta < x_{(2)} \\ & \vdots \\ \sum_{i=j+1}^{n} x_{(i)} - \sum_{i=1}^{j} x_{(i)} + (2j-n)\theta & x_{(j)} \le \theta < x_{(j+1)} \\ & \vdots \\ x_{(n)} - \sum_{i=1}^{n-1} x_{(i)} + (n-2)\theta & x_{(n-1)} \le \theta < x_{(n)} \\ -\sum_{i=1}^{n} x_{(i)} + n\theta & x_{(n)} \le \theta \end{cases}$$

Provided $2j - n > 0$, this is an increasing function.

If n is odd, then the function peaks at $X_{(j+1)}$.
If n is even, the function is flat $(2j-n=0)$ over $\theta \in (X_{(j)}, X_{(j+1)})$.

So the MLE is

$$\hat{\theta}_n \in \begin{cases} \{X(\lceil \frac{n}{2} \rceil)\} & \text{if } n \text{ is odd} \\ [X(\frac{n}{2}), X(\frac{n}{2}+1)] & \text{if } n \text{ is even.} \end{cases}$$

$\boxed{7.14}$ From exercise 4.26, we found

$$Z \sim f_Z(z) = \left(\frac{\lambda+\mu}{\lambda\mu}\right) e^{-z\left(\frac{\lambda+\mu}{\lambda\mu}\right)} \mathbb{1}(z > \mu)$$

$$W \sim \text{Bernoulli}\left(\frac{\mu}{\lambda+\mu}\right)$$

with $Z \perp\!\!\!\perp W$.

Let $(Z_1, W_1), \ldots, (Z_n, W_n)$ be independent realizations of (Z, W).

Since W is discrete and Z is continuous, the likelihood is somewhat awkward, but we can write it down as

$$L(\lambda, \mu; \underset{\sim}{Z}, \underset{\sim}{W}) = \prod_{i=1}^{n} \left(\frac{\lambda+\mu}{\lambda\mu}\right) e^{-z_i\left(\frac{\lambda+\mu}{\lambda\mu}\right)} \cdot \left(\frac{\mu}{\lambda+\mu}\right)^{w_i} \left(\frac{\lambda}{\lambda+\mu}\right)^{1-w_i}$$

$$= \left(\frac{\lambda+\mu}{\lambda\mu}\right)^{n} e^{-n\bar{Z}_n\left(\frac{\lambda+\mu}{\lambda\mu}\right)} \cdot \left(\frac{\mu}{\lambda+\mu}\right)^{n\bar{W}_n} \left(\frac{\lambda}{\lambda+\mu}\right)^{n-n\bar{W}_n}$$

and the log-likelihood is

$$\ell(\lambda, \mu; \underset{\sim}{z}, \underset{\sim}{w}) = n \log(\lambda+\mu) - n \log \lambda - n \log \mu - n \bar{z}_n \left(\frac{\lambda+\mu}{\lambda\mu}\right)$$

$$+ n \bar{w}_n \log \mu - n \bar{w}_n \log(\lambda+\mu)$$

$$+ (n - n\bar{w}_n) \log \lambda - (n - n\bar{w}_n) \log(\lambda+\mu) .$$

Now

$$\frac{\partial}{\partial \mu} \ell(\lambda, \mu; \underset{\sim}{z}, \underset{\sim}{w}) = \frac{n}{\lambda+\mu} - \frac{n}{\mu} + \frac{n\bar{z}}{\mu^2} + \frac{n\bar{w}_n}{\mu} - \frac{n\bar{w}_n}{\lambda+\mu} - \frac{(n-n\bar{w}_n)}{\lambda+\mu}$$

$$= \frac{n\bar{w}_n - n}{\mu} + \frac{n\bar{z}}{\mu^2} \overset{\text{set}}{=} 0$$

$$\Rightarrow \mu = \frac{\bar{z}}{1-\bar{w}_n}$$

$$\frac{\partial}{\partial \lambda} \ell(\lambda, \mu; \underset{\sim}{z}, \underset{\sim}{w}) = \frac{n}{\lambda+\mu} - \frac{n}{\lambda} + \frac{n\bar{z}_n}{\lambda^2} - \frac{n\bar{w}_n}{\lambda+\mu} + \frac{n-n\bar{w}_n}{\lambda} - \frac{n-n\bar{w}_n}{\lambda+\mu}$$

$$= -\frac{n\bar{w}_n}{\lambda} + \frac{n\bar{z}_n}{\lambda^2} \overset{\text{set}}{=} 0$$

$$\Rightarrow \lambda = \frac{\bar{z}_n}{\bar{w}_n} .$$

So we have

$$(\hat{\mu}, \hat{\lambda}) = \left(\frac{\bar{z}_n}{1-\bar{w}_n}, \frac{\bar{z}_n}{\bar{w}_n}\right)$$

$\boxed{7.15}$ Let $X_1, \dots, X_n \overset{ind}{\sim} f(x; \lambda, \mu) = \sqrt{\dfrac{\lambda}{2\pi x^3}} \exp\left[-\lambda(x-\mu)^2/(2\mu^2 x)\right], \ x>0$

The likelihood function is given by

$$L(\lambda, \mu; \underset{\sim}{X}) = \prod_{i=1}^{n} \sqrt{\frac{\lambda}{2\pi X_i^3}} \exp\left[-\lambda (X_i-\mu)^2/(2\mu^2 X_i)\right]$$

$$= (2\pi)^{-n/2} \lambda^{n/2} \left(\prod_{i=1}^{n} X_i\right)^{3/2} \exp\left[-\lambda \sum_{i=1}^{n} \frac{(X_i-\mu)^2}{2\mu^2 X_i}\right]$$

and the log-likelihood is

$$\ell(\lambda, \mu; \underset{\sim}{X}) = -\frac{n}{2}\log(2\pi) + \frac{n}{2}\log\lambda + \frac{3}{2}\sum_{i=1}^{n}\log X_i - \frac{\lambda}{2\mu^2}\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{X_i}.$$

We have

$$\frac{\partial}{\partial\lambda}\ell(\lambda, \mu; \underset{\sim}{X}) = \frac{n}{2\lambda} - \frac{1}{2\mu^2}\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{X_i} \overset{set}{=} 0$$

$$\frac{\partial}{\partial\mu}\ell(\lambda, \mu; \underset{\sim}{X}) = \frac{\lambda}{\mu^3}\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{X_i} + \frac{\lambda}{\mu^2}\sum_{i=1}^{n}\frac{(X_i-\mu)}{X_i} \overset{set}{=} 0$$

The first equation gives $\lambda = \dfrac{n\mu^2}{\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{X_i}}$.

Plugging this into the second equation gives

$$\frac{n}{\mu^3} + \frac{n}{\mu^2}\frac{1}{\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{X_i}}\sum_{i=1}^{n}\frac{(X_i-\mu)}{X_i} \overset{set}{=} 0$$

Now write

$$\sum_{i=1}^{n}\frac{(X_i-\mu)}{X_i} = \sum_{i=1}^{n}\left(1 - \frac{\mu}{X_i}\right) = n - \mu\sum_{i=1}^{n}\frac{1}{X_i}$$

$$\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{X_i} = \sum_{i=1}^{n}\frac{X_i^2 - 2X_i\mu + \mu^2}{X_i} = \sum_{i=1}^{n}X_i - 2n\mu + \mu^2\sum_{i=1}^{n}\frac{1}{X_i}$$

so that

$$\frac{n}{\mu^3} + \frac{n}{\mu^2}\,\frac{n - \mu \sum_{i=1}^{n} \frac{1}{x_i}}{\sum_{i=1}^{n} x_i - 2n\mu + \mu^2 \sum_{i=1}^{n} \frac{1}{x_i}} = 0$$

If we plug in $\bar{x}_n$ for μ, we obtain

$$\frac{n}{\bar{x}_n^3} + \frac{n}{\bar{x}_n^2}\,\frac{n - \bar{x}_n \sum_{i=1}^{n} \frac{1}{x_i}}{\sum x_i - 2n\bar{x}_n + \bar{x}_n^2 \sum_{i=1}^{n} \frac{1}{x_i}}$$

$$= \frac{n}{\bar{x}_n^3} + \frac{n}{\bar{x}_n^2}\,\frac{n - \bar{x}_n \sum_{i=1}^{n} \frac{1}{x_i}}{n\bar{x}_n - 2n\bar{x}_n + \bar{x}_n^2 \sum \frac{1}{x_i}}$$

$$= \frac{n}{\bar{x}_n^3} + \frac{n}{\bar{x}_n^3}\left(\frac{n\bar{x}_n - \bar{x}_n^2 \sum_{i=1}^{n} \frac{1}{x_i}}{-n\bar{x}_n + \bar{x}_n^2 \sum_{i=1}^{n} \frac{1}{x_i}}\right)$$

$$= \frac{n}{\bar{x}_n^3} - \frac{n}{\bar{x}_n^3}$$

$$= 0$$

So $\hat{\mu}_{mle} = \bar{x}_n$. Plugging this into the expression for λ gives

$$\hat{\lambda}_{mle} = \frac{n\bar{x}_n^2}{\sum_{i=1}^{n} \frac{(x_i - \bar{x}_n)^2}{x_i}}$$

$$= \frac{n\bar{x}_n^2}{\sum_{i=1}^{n}\left(\frac{x_i^2 - 2x_i\bar{x}_n + \bar{x}_n^2}{x_i}\right)}$$

$$= \frac{n\bar{x}_n^2}{\sum_{i=1}^{n} x_i - 2n\bar{x}_n + \bar{x}_n^2 \sum_{i=1}^{n} \frac{1}{x_i}}$$

$$= \frac{n\bar{x}_n^2}{-n\bar{x}_n + \bar{x}_n^2 \sum_{i=1}^{n} \frac{1}{x_i}}$$

$$= \frac{n}{\sum_{i=1}^{n}\left(\frac{1}{x_i} - \frac{1}{\bar{x}_n}\right)}.$$